



DEUTSCHE INITIATIVE  
FÜR NETZWERKINFORMATION E.V.

## **Standardised Usage Statistics for Open Access Repositories and Publication Services**

DFG Project "Open Access Statistics"  
and  
DINI Working Group "Electronic Publishing"





**DEUTSCHE INITIATIVE  
FÜR NETZWERKINFORMATION E.V.**

**Standardised Usage Statistics  
for Open Access Repositories  
and Publication Services**

DFG Project "Open Access Statistics"  
and  
DINI Working Group "Electronic Publishing"

## **Imprint**

DINI – Deutsche Initiative für Netzwerkinformation e. V.

DINI-Geschäftsstelle

c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen

Platz der Göttinger Sieben 1

D-37073 Göttingen

Tel.: +49 551 39-33857

Fax: +49 551 39-5222

E-Mail: [gs@dini.de](mailto:gs@dini.de)

[www.dini.de](http://www.dini.de)

# Contents

	<b>About DINI</b>	<b>5</b>
	<b>Summary</b>	<b>6</b>
<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Open Access statistics</b>	<b>9</b>
<b>3</b>	<b>Standards and protocols</b>	<b>11</b>
3.1	Evaluation	12
3.2	Survey results	13
3.3	Conclusion	16
<b>4</b>	<b>Implementation</b>	<b>19</b>
4.1	Processing access information	20
4.2	Returning aggregated usage figures to the repositories	22
4.3	Optimising the technology and consolidating usage figures	23
	Excursus: Report from pilot partner EconStor	23
4.4	Data protection	25
4.5	Filtering non-human hits	26
<b>5</b>	<b>Current developments and requirements</b>	<b>28</b>
<b>6</b>	<b>Bibliography</b>	<b>33</b>
	<b>Index of authors</b>	<b>35</b>



## About DINI

The development of modern information and communication technologies causes a change in the information infrastructures of higher education institutions and other research institutions. This change is a major topic within higher education in Germany, and more than ever requires agreements, cooperation, recommendations, and standards. The Deutsche Initiative für Netzwerkinformation (DINI, German Initiative for Network Information) supports this development.

DINI was founded to advance the improvement of the information and communication services and the necessary development of the information infrastructures at the universities as well as on regional and national levels. Agreements and the distribution of tasks among the infrastructure facilities can significantly extend the range of information technology and of services. Additionally, the joint development of standards and recommendations is a requirement.

DINI is an initiative of three organizations:

- AMH (Arbeitsgemeinschaft der Medienzentren der deutschen Hochschulen; Consortium of German University Media Centers),
- dbv (Deutscher Bibliotheksverband Sektion 4: Wissenschaftliche Universalbibliotheken; German Library Association, Section 4: Academic Universal Libraries),
- ZKI (Zentren für Kommunikation und Informationsverarbeitung in Lehre und Forschung e. V.; Association of German University Computing Centers).

DINI has the following goals:

- Publicize and recommend best practices;
- Encourage and support the formulation, application and further development of standards as well as distribute recommendations regarding their application;
- Register and advertise Competence Centers using modern web-based technologies;
- Improve inter-disciplinary exchange through congresses, workshops, expert conferences etc.;
- Advertise new funding programs and encourage new programs.

## Summary

The Open Access statistics (OA statistics) service provides repositories and publications services with standardised usage statistics for documentation. Its declared aim is to increase acceptance of Open Access among authors and readers of scientific publications. In contrast to citation-based metrics which retrospectively measure the impact of a publication, usage statistics for a digital document provide that particular document's current level of relevance, thus also mapping dynamic online trends.

An infrastructure was created within the scope of the DFG-funded OA statistics project in order to collect and process standardised usage data to be passed on to the head office of the GBV Common Library Network following completion of the second funding phase in summer 2013.



# 1 Introduction

The reputation of a scientist within the academic world is based on the importance of his or her work and scientific impact. The impact of this person's scientific publications is traditionally linked to the citation-based Journal Impact Factor (JIF). A higher Journal Impact Factor is universally taken to be a sign of the quality of a journal and its articles. This value is of particular importance in natural sciences and medicine where it is seen as a kind of quality seal for scientific performance. A scientist's career often depends on how many of his or her articles appear in journals with a high JIF.

The JIF is a measure reflecting the average number of citations to articles published in journals. This therefore shows how often any of the journal's articles are cited on average within a certain period of time (usually two years) prior to the reference year for calculating the JIF, rather than how often one specific article is cited. The bases for calculating this value are the citations from all of the journals in the Web of Science, a bibliographical database that can be licensed for a fee. The Web of Science largely involves international journals that only publish articles that have been successfully peer reviewed. The JIF therefore receives a fair amount of criticism<sup>1</sup> as this citation frequency measurement does not by any means cover all scientific publications. Certain document types are completely ignored, and calculations are based on entire journals rather than individual articles.

In the era of digital online publications, fewer articles are being published in print media. Download figures can now be established as usage statistics which represent an alternative to JIF as well as usage impact. These statistics can be collected in real time with little effort and analysis taking place at object level (cf. table 1). This allows a usage impact value to be determined for each article. Citation-based metrics, on the other hand, can only be measured with a delay as the citations are counted retrospectively, i.e. after the cited article has been published, and it is this figure that is used to calculate the impact factor. On top of that, only the journals listed in the index are included in the calculation. Usage, on the other hand, can be measured immediately following publication with the counter increasing after every hit. When it comes to citation-based metrics, it is the authors who determine the metric by referencing publications within the scope of scientific correspondence. With user-based metrics, on the other hand, it is

---

<sup>1</sup> Cf. Dong (2005) and Seglen (1997).

the readers who determine the metric by clicking on the individual documents of interest to them. This means that usage statistics allow us to make statements regarding the temporal distribution and usage of publications.

Citation-based metrics		Usage-based metrics
Author	Who?	Reader
Delayed	When?	Immediately
Indexed journals	What?	All digital objects
Analysis at journal level	How?	Analysis at object level

*Table 1: Differences between citation-based and usage-based metrics*

Open Access publications are not subject to any access restrictions, meaning that they are an excellent basis for collecting usage statistics and for establishing an internationally comparable standard. This in turn means that as a form of publication, Open Access represents a low-cost and transparent way to evaluate the distribution and usage of scientific findings. The OA statistics project embraced these ideas in order to boost the acceptance of Open Access among authors and recipients of scientific publications by generating comparable usage statistics and offering a lasting infrastructure to collect and process usage data. The OA statistics service provides aggregated and standardised usage data for electronic scientific documents made available in repositories according to the Open Access principle. Thanks to the OA statistics service, these documents can be provided with usage-based evaluations.

## 2 Open Access statistics

From 2008 onwards, the OA statistics project also comprised a service aimed at calculating and providing comparable standardised usage statistics for scientific Open Access repositories and the journals and documents published there. In order to provide standardised usage statistics, OA statistics created an infrastructure to exchange and aggregate usage data. The project partners as of the first funding phase were the Göttingen State and University Library, the computer and media service at the Humboldt-Universität zu Berlin, the Saarland University and State Library, and the Stuttgart University Library. Another project partner – the head office of the GBV Common Library Network – was joined the second phase of the project between 2011 and 2013. In order to ensure internationally agreed standards for exchanging and calculating usage data, the project partners cooperated with partners in Germany and other countries.<sup>2</sup> Together with the usage statistics knowledge exchange working group,<sup>3</sup> OA statistics put together a series of guidelines<sup>4</sup> to standardise the exchange of usage data on a European level. The first phase of this project ended in December 2010, by which time a basis for a lasting infrastructure to collect and process usage data had been created. The focus of this infrastructure was to develop and establish a uniform standard to determine the number of hits and offer added-value services to repositories. Due to the project members' involvement in the DINI Electronic Publishing working group, the OA statistics results were incorporated into the DINI 2010 certificate<sup>5</sup> which explicitly recommends the OA statistics infrastructure for certified repositories. The second phase of the project was launched in April 2011. Here the focus was on expanding the infrastructure to include more German repositories and swiftly and reliably provide standardised usage statistics. Funding of the project ended in summer 2013, after which time the head office of the GBV Common Library Network took over the service and added it to its portfolio. This involves both technical operation of the service on the head office of the GBV Common Library Network's servers and handling of the entire financial and contractual situation. The head office of the GBV Common Library Network is supported by

---

<sup>2</sup> Partners include OA-Netzwerk (Germany), COUNTER (Great Britain), ROAT (Japan) and OpenAIRE (Europe).

<sup>3</sup> <http://www.knowledge-exchange.info/Default.aspx?ID=365>.

<sup>4</sup> Cf. Verhaar (2011).

<sup>5</sup> <http://www.dini.de/dini-zertifikat/>.

the DINI Electronic Publishing working group and personal contributions from the project partners. The feasibility study<sup>6</sup> carried out during the project term and the project's data protection law review<sup>7</sup> were the reasons for turning the project into a service. Anyone in charge of repositories who is interested in this service can contact the head office of the GBV Common Library Network or any of the former OA statistics project partners. After paying a small fee and submitting usage data, interested parties will then be sent usage statistics that have been aggregated and are in line with international standards.

---

<sup>6</sup> Cf. OA statistics (2012).

<sup>7</sup> Cf. ZENDAS (2011).

### 3 Standards and protocols

The OA statistics project work involved collaborations with a number of different institutions and projects from various countries so that uniform standards are used whenever possible to facilitate an exchange of usage data. Prior to the first funding phase, a JISC workshop<sup>8</sup> was held where the participants agreed to use OpenURL ContextObjects (NISO/ANSI 39.88-2004 standard)<sup>9</sup> and OAI-PMH as standards.<sup>10</sup> Within the scope of the usage statistics knowledge exchange working group, the project partners also put together a series of guidelines to standardise the exchange of usage data.

Following this exchange, the usage data had to be aggregated using internationally agreed processes. Providers of scientific literature often complain about the lack of consistency in analysing bibliometric data, which is why a number of experts were interviewed about conventional standards during the OA statistics project. The evaluation of conventional standards included the Counting Online Usage of Networked Electronic Resources (COUNTER)<sup>11</sup> method used by libraries as an established industry standard for journal evaluations, the LogEc<sup>12</sup> method used to evaluate the economic database Research Papers in Economics (RePEc), and the IFABC<sup>13</sup> method to measure hit frequencies in the advertising industry. The aim of this evaluation was to gather practical information about these methods and then decide which one was best for use with OA statistics. Another aim was to generate knowledge about the individual key figures produced by the respective method. The idea behind this was to isolate the individual pros and cons rather than the method as a whole in order to be able to make suggestions for new methods or to support the modification of existing standards. In addition, the OA statistics project was also designed to uncover any additional requests that could then be used to develop a needs-based service.

---

<sup>8</sup> Cf. Merk (2008).

<sup>9</sup> [http://www.niso.org/apps/group\\_public/project/details.php?project\\_id=82](http://www.niso.org/apps/group_public/project/details.php?project_id=82).

<sup>10</sup> Cf. Bollen (2006).

<sup>11</sup> <http://www.projectcounter.org/>.

<sup>12</sup> <http://logec.repec.org/>.

<sup>13</sup> <http://www.auditbureau.org.au/ifabctest/>.

### 3.1 Evaluation

The OA statistics project not only involved setting up the service, it also included an analysis of conventional methods used to calculate usage standards and a discussion to determine the need for any additional features. To this end, the project plan included two evaluations.<sup>14</sup> A survey conducted during the first phase of the project<sup>15</sup> was used to draw a comparison of how the situation has developed over the last few years. The main questions in both evaluations can be summarised as follows:

- To what extent can Open Access documents convey prestige and activity by using various metrics, and which of the available standards are most suitable in practice? In order to answer this question, an analysis was carried out involving calculation methods deemed to be standards like COUNTER, LogEc and IFABC. The focus here was on the parameters that the respective methods use to define their characteristic values.
- Aside from pure usage statistics, which additional features are of interest when it comes to Open Access publication offerings? The survey conducted during the first phase of the project was modified and then carried out again among repository specialists. The idea behind this was to rank potential additional features that needed to be implemented during this phase of the project.

In order to perform this analysis, a total of 32 specialists in Germany and other countries were contacted and invited to take part in the two surveys. Eight people completed the first survey, which equated to an above-average response rate of 25%. The second survey was completed in full by nine people, which meant a slightly higher response rate of 28%.

The method used in both surveys was based on a combination of standardised and open questions in the form of expert interviews conducted using online survey software. The surveys consisted of 97 questions split up into seven sections (personal details, opinions, usage, information, general and direct comparison, problems, requirements, requests and outlook). The surveys were not anonymous so any follow-up questions could be clarified in the form of telephone interviews. As a result, the surveys used a method mix that provided a comparatively high density of information.

---

<sup>14</sup> The full report is available here: [http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/OAS\\_Bericht\\_Evaluation.pdf](http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/OAS_Bericht_Evaluation.pdf)

<sup>15</sup> Cf. Herb (2012).

### 3.2 Survey results

The aim of the first survey was to evaluate currently established standards. The general conclusion of this survey was that COUNTER was considered to be the most important standard by the survey participants. Almost 85% of the people surveyed were either happy or very happy with COUNTER (cf. table 2). A number of different reasons were given for this positive opinion, although most were coupled with requests for improvement: "They (COUNTER) provide useful information at journal level. The standard would be even more useful than it already is if it also included information on items (as proposed in the PIRUS project)." The reasons for this largely positive view of COUNTER ranged from general statements such as "COUNTER is the most common standard" and "globally recognised" through to more detailed explanations such as the "exchange of usage data for certain items, with some descriptive metadata". Some people did however also add that it will only be possible to analyse metadata at article level following completion of the implementation phase at the end of 2013. The survey participants also criticised COUNTER's lack of separation when it comes to collecting data (at article level) and presenting data (aggregated at journal level without temporal differentiation), a robot list whose creation lacks transparency and is largely non-systematic, a double-click interval that is too short, the lack of an option to compare conventional and Open Access publications, and insufficient report documentation: "A lot of information is missing in the resulting reports: no user identification, only minimal metadata about the item, no referrer info." Despite this criticism, most of the responses were positive with many of the participants stating that COUNTER produces sound statistics thanks to its standardisations, enables comparisons that were previously not possible, and also acts as a good basis for future analyses: "Being a recognised standard, it is a good basis for underpinning other things such as altmetrics which can then indicate the context of use."

Despite the generally positive opinions of COUNTER, it is still not considered the best standard when directly comparing its individual components with those of other standards. This therefore gives rise to the hypothesis that this discrepancy is due to assumptions and stereotypes having a much stronger impact the more general the rating level (i.e. first impression). Assumptions have an increasingly diminishing role to play the more detailed the rating. Real knowledge is required here, which, among other things, leads to contrasting results. In order to establish

a standard for usage analyses, it appears to be a good idea to adopt a proactive approach to advertising such a standard, thus drawing as much attention to it as possible.

Do you agree that COUNTER/LogEc/IFABC is a suitable standard for your work?			
	COUNTER	LogEc	IFABC
Strongly disagree	7,7%	7,7%	7,7%
Somewhat disagree	7,7%	15,4%	
Don't know		15,4%	15,4%
Somewhat agree	53,8%	23,0%	15,4%
Strongly agree	30,8%	7,7%	
Not familiar with ...		30,8%	61,5%

Table 2: Suitable standard for work

Please rate several COUNTER criteria by comparing it to LogEc <sup>16</sup> . Do you generally consider COUNTER's criteria to be much worse, worse, as good as, better, or much better than LogEc's?						
	Counter criteria COUNTER vs. LogEc	Multiclick interval COUNTER vs. LogEc	User identification COUNTER vs. LogEc	Crawler Definition COUNTER vs. LogEc	Crawler identification COUNTER vs. LogEc	Crawler counter COUNTER vs. LogEc
Much worse	0%	0%	0%	0%	0%	0%
Worse	20%	40%	20%	20%	40%	0%
As good as	40%	20%	20%	20%	0%	20%
Better	20%	0%	20%	20%	20%	20%
Much better	0%	0%	0%	0%	0%	0%
Don't know	20%	40%	40%	40%	40%	60%

Table 3: COUNTER vs. LogEc

<sup>16</sup> The IFABC standard is missing from this comparison as almost two thirds of the people surveyed stated that they had absolutely no prior experience with it. Most of the people who stated they had some IFABC experience said that this was not enough to be able to compare it with COUNTER.



From an expert point of view, COUNTER is the best solution and technical shortcomings are considered to be of minor importance. Interestingly enough, the experts appeared to be swayed more by subjective presumptions than detailed facts. People's decisions are therefore driven by the standard's image within the corresponding communities rather than its technical capabilities. This also reflects the need and desire for standardisation. Nowadays, there is practically no way around COUNTER for anyone looking to portray their service as being efficient, serious, reliable and sustainable as COUNTER was not only considered the best standard, it was also deemed to be globally recognised and therefore the most common standard. COUNTER's significance as a standard can be seen in the fact that it was used as a blueprint for contemplating a fictive "own" ideal calculation standard.

Irrespective of the real requirements set out by COUNTER, LogEc and IFABC, the people surveyed were asked to draft and outline their ideal standard. This ideal standard presents results in JSON in the form of a detailed visualisation. Not only that, it also processes information quickly, is considered state of the art and well-known among the various respective communities, has a constantly updated robot list, effective access cleaning (in particular with regard to double clicks), is intuitive and therefore easy to set up and use. The standard is also expected to be future-proof and flexible, with flexibility applying in particular to the major changes taking place in terms of media visualisation while taking account of the requirements arising from the use of social media. One expert put it as follows: "1. Fast 2. Simple 3. Extendable 4. Easily interoperable with other services."

The repeated reference to alternative media types underlines the growing importance of social media in science. Every survey participant acknowledged this, although there are a number of biases against new publication types in the somewhat conservative field of science of which the participants were also aware. The point here is to introduce speed, spread knowledge and implement informally established methods in scientific communication, which is why the people surveyed think social media have a major role to play. Despite this, the survey participants did not return a clear opinion as to whether or not social media could help to generate scientific recognition. A somewhat more positive opinion was received in response to the question of whether or not social media could improve impact both within scientific circles and beyond. 60% of those surveyed said yes and 40% said no to the former question, with 70% saying yes and 30% saying no to the latter. 100% of the experts said yes to the question of whether social media could contribute to an author's reputation within the world of science. It would

seem that a clear distinction is made between scientific acclaim and personal reputation, as is the case with scientific and non-scientific domains. As a result, social media are deemed to be insufficient for imparting complex scientific results, but well-suited to portraying and profiling the author of these complex scientific results. Social media are therefore great for presentation purposes, but do not provide any information about the actual product itself. This in turn means that a mix of methods is required to achieve maximum impact.

Social media are currently considered to lack content with experts preferring more conventional forms of digital and analogue publications. This indicates the importance of verifying the quality of a scientific article. A number of survey participants had reservations and ideas in terms of reorganising scientific ratings:

*I found the Priem / Hemminger paper<sup>17</sup> on the 'decoupled journal' very inspiring. I guess it's better not to have every single journal organise their 'own' peer review process. P. review is produced by the scientific community (for free!), anyway, so why not find better, more transparent, maybe faster ways? I'm sure we live in a 'first publish, then filter' world (Clay Shirky)<sup>18</sup> anyway. I don't see much use in a closed process of review that has to happen all the way BEFORE publication happens. It has to be easy to pick up on any 'finished' result of scientific research anyway. Criticism of results won't be separable from open, reproducible research results in the long run. Reproducibility and assessability are two sides of the same coin.*

In times of sites like Wikipedia, perhaps this 'first publish, then filter' approach is the perfect reaction in terms of restructuring (not just) science. This (scientific) impact is the result of voting by mouse click, which in turn means that voting results can be visualised and made tangible in the sense of a public vote, thus providing retrospective legitimisation: "The future presented by the internet is the mass amateurisation of publishing and a switch from 'Why publish this?' to 'Why not?'"<sup>19</sup>.

### 3.3 Conclusion

The standards survey conducted among experts returned a clear general consensus for a uniform and therefore comparative standard. There was also a tendency to want something that was technically feasible, but this came about with practi-

---

<sup>17</sup> Cf. Priem (2012).

<sup>18</sup> Cf. Shirky (2008).

<sup>19</sup> Cf. *ibid.*, p. 60.

cally no consideration whatsoever for aspects such as data protection or financing. The survey showed that COUNTER was the clear leader of the pack, but trailed its rival LogEc (cf. table 3) in certain areas such as double-clicks interval and crawler identification.

COUNTER defines uniform standards as to how information usage can be measured. To this end, COUNTER codes of practice<sup>20</sup> were developed which OA statistics use as a standard for aggregating usage data and for calculating usage statistics. Within a library context, the aggregated COUNTER figures are primarily used to determine how publications licensed by publishing houses are used in the respective library or consortium. Publications providers also benefit from this standard as it gives them access to consistent usage statistics, which they can in turn use as a sales argument. COUNTER uses a robot blacklist to eliminate automated hits. Such a list is of far greater importance when using the COUNTER standard with OA repositories as opposed to a conventional library context as Open Access documents are subject to far more machine hits than publications to which access is blocked. Such a list needs to be continually coordinated and expanded to include new robots. Usage statistics can only be comparable on an international level if every repository uses the same list.

The results of the importance of the various additional usage statistics features showed that the people surveyed set great store on networking and collaboration with other media and social media as well as recommendation features: "Rating is not always as important as sharing or bookmarking." Pure additional information and services are definitely seen to represent added value, but are of lesser importance when compared directly. The survey produced the following overall rating: 1) Networking, 2) Recommendations, 3) Information, 4) Service. It should however be pointed out that the results of the survey also partially contradicted one another the more specific the questions were.

When directly comparing this survey with the one carried out in 2009 it can be seen that networking with other media and social media has risen significantly in terms of importance. In 2009 networking was not considered important at all, yet in 2012 it was considered vital. In contrast to that, the creation of recommenders and rankings as a form of orientation for users was deemed unimportant. The fact that the networking and recommendation features were considered very important is of little surprise as they are mutually dependent. The expansion of networking to include other media types and social media help to boost impact

---

<sup>20</sup> [http://www.projectcounter.org/code\\_practice.html](http://www.projectcounter.org/code_practice.html).

and range, which in turn increases the complexity of the usage scope. Here, recommendations are far more than just filters and a form of orientation. The pure production of information and provision of various services correspond to the old analogue library world and are considered representatives of the “old system”. On the other hand, the categories “Networking” and “Recommendations” are seen as the future of scientific publishing and literature management.

The most added value can be seen in the opportunity to network between documents in repositories and documents located elsewhere (journals, electronic platforms) and the opportunity to link repositories with password-protected social networks for scientists. Although current trends are showing a shift towards social media, experts still consider the analysis of hits to be the most important aspect. The experts therefore gave due consideration to the status quo but answered with a view to the future, most likely having taken significant events into account.

To summarise, opinions are clearly split when it comes to additional features with networking, collaboration with other media and recommendations deemed to be most important, while the actual information and services are considered to be relatively unimportant.

## 4 Implementation

A technical infrastructure is required to collect, exchange, aggregate and provide usage statistics for digital objects. The OA statistics infrastructure consists of two core components: The OA statistics data provider (OAS data provider), which has to be provided by the repository operator, is installed in the respective repository. The installation software can be downloaded<sup>21</sup> from the SourceForge website, but must then be adapted to the specific repository requirements. The service provider uses the OAS service provider which collects and aggregates the data from the OAS data providers and then supplies the standardised usage data.

Repositories wanting to offer their users internationally standardised usage figures can get involved in the OAS infrastructure as an OAS data provider. Help is available in the form of the installation guide description<sup>22</sup>, the OAS data provider demo installation<sup>23</sup> and the OAS validator<sup>24</sup>. In the event of queries which none of the above documents can answer, users can contact the OA statistics support<sup>25</sup> for further assistance. The latest version of the OAS data provider software is available for various repository platforms, but must be adapted and configured to meet local requirements. The demo installation provides a number of examples on how to return aggregated usage statistics to XML and JSON formats as well as an example of graphical integration using iFrame. After installing a new OAS data provider, the data format output by the interface can be checked using the OAS validator. If the XML contained there is valid, the OAS data provider can register using the registration form<sup>26</sup>. OA statistics then contacts the contact person of the respective repository in order to clarify the return format of the aggregated data and the exchange of passwords.

---

<sup>21</sup> <https://sourceforge.net/p/openaccessstats/code-0/HEAD/tree/trunk/>.

<sup>22</sup> [http://www.dini.de/fileadmin/oa-statistik/technik/Leitfaden\\_fuer\\_neue\\_Repositorien\\_v1.0.pdf](http://www.dini.de/fileadmin/oa-statistik/technik/Leitfaden_fuer_neue_Repositorien_v1.0.pdf).

<sup>23</sup> <http://oas-cdp.gbv.de/oai-data-provider.demo/index.php?verb=ListRecords&metadataPrefix=oas>.

<sup>24</sup> <http://oas-sp.gbv.de/validator/>.

<sup>25</sup> <http://www.dini.de/projekte/oa-statistik/kontakt/>.

<sup>26</sup> <http://www.dini.de/projekte/oa-statistik/teilnahme-am-dienst/registrierungsformular/>.

## 4.1 Processing access information

In OA statistics, usage statistics for Open Access documents are calculated by standardising the number of document hits and evaluating them centrally. Below is an explanation of how statistics are generated from document hits (cf. figure 1).

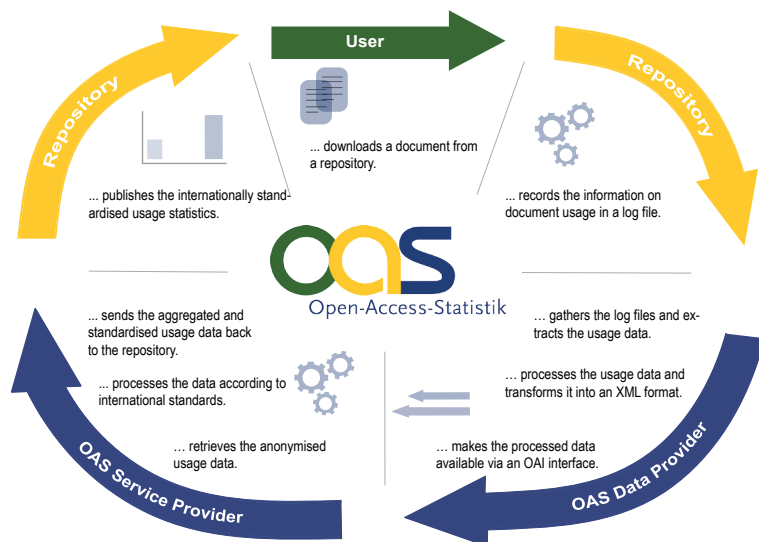


Figure 1: OA statistics workflow

### A repository user downloads a document from the repository

If repository users request a document, their browser automatically transfers data such as their user agent (incl. OS and browser version), ACCEPT header, preferred languages, protocol version, IP address, and name of the requested document.

### Repository saves the usage information in log files

An entry is made in the web server's log files every time someone accesses a repository's documents. For reasons pertaining to data protection and/or media law, the repository provides an opt-out option. If the user opts out, their IP address is not stored in the log files and therefore cannot be processed.

### OAS data provider collects usage information from the log files and pseudonymises the user ID

Whether or not IP addresses constitute personal data is still a major topic of discussion in Germany. The prevailing view among data protection officers there is that this information constitutes personal data, and the European Court of Justice

has also taken up this stance (case C-70/10). Nevertheless, a user's IP address is indisputably personal when it comes to repository operators who are simultaneously access providers. This is why IP addresses added to log files when using documents are SHA-256 hashed and salted as this pseudonymises the data while allowing for the identification of a usage pattern of a certain IP address without actually processing the IP address itself. Salting is another protective measure used to prevent the hash value from being retraced.

The C class network of the IP address (the first three digits of an IP address) is also salted and hashed to take account of subsequent robot hits and to be able to analyse the data.

### **OAS data provider processes the usage information and converts it into an XML format containing clear document identifiers**

The OAS data provider's web server log files are converted into Open URL ContextObjects in the log file parser before being saved in XML format.

An identifier for the respective document also has to be included along with the information from the log files. In most cases, this takes place based on the invoked URL or the repository's individual ID number. This process is different for every repository, meaning that it has to be adapted during installation in the OAS data provider's log file parser.

The OpenURL ContextObjects format was adapted by OA statistics and then matched to other European projects before being documented.<sup>27</sup> This is more complex than the minimum set of OpenURL ContextObjects, but is still fully compatible.

### **OAS data provider provides usage information via OAI interface**

The OAS data provider reads the OpenURL ContextObjects from a database and provides it to the OAS service provider via the OAI protocol. The OAS data provider is an independent program that differs from other OAI data providers and has to be installed separately.

### **OAS service provider accesses anonymised usage information via OAI interface**

The OAS service provider retrieves the data from the OAS data provider using an OAI interface. For data protection reasons, this data must then be deleted from the OAS data provider.

---

<sup>27</sup> The data format specification is available by visiting [http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/Specification\\_V5.pdf](http://www.dini.de/fileadmin/oa-statistik/projektergebnisse/Specification_V5.pdf).

## **OAS service provider processes usage information according to international standards**

The usage information is analysed according to the COUNTER standard by filtering out double clicks by users as well as robot hits.

## **OAS service provider supplies the aggregated and standardised usage data**

The usage data is then provided to the repository and other interested parties based on the licencing requirements. The interface and format are agreed on in advance with the repository operators.

## **Repository provides access to internationally standardised usage statistics**

Daily updated usage figures for individual documents are supplied to the participating repositories. This may comprise entire extracts of the OAS service provider's database available via an interface or an iFrame provided by OA statistics that can be integrated into the repository. In this case the document identifier also has to be supplied as a parameter.

### **4.2 Returning aggregated usage figures to the repositories**

The OAS service provider returns the aggregated usage figures for the participating repositories in two forms. Each repository initially receives password-protected access to an Apache web server containing usage statistics files stored and updated at defined intervals. The scope of each individual file, evaluation frequency and data format are agreed on in advance with the repository. An index file listing the most recently created files serves to simplify their automated retrieval and also indicates any updates to the data.

The second form is where OAS data providers can automatically retrieve usage figures via a HTTP-based REST interface. The scope of data involved can be specified, e.g.:

- The desired period of time
- The format
- The identifier, whereby data can be output for a specific or all identifiers
- The desired categories (full-text COUNTER, abstract COUNTER, full-text robot hits, abstract robot hits)

The output data also contains a list of the usage data per document identifier extrapolated to a day, a week, a month or a year. Daily output is most common and most precise. Data output can therefore be specified for any particular period of time and various exchange formats are available for selection in agreement with the participating repositories. The standard formats are JSON and CSV,



although two XML formats based on the version 4 COUNTER standard are also available. Tabular calculation programs are also catered for thanks to generic XML and SpreadsheetML-based versions. Both of these versions can in turn be used to create a COUNTER standard-specific Journal Report 1 (JR1) and Book Report (BR1).<sup>28</sup>

### 4.3 Optimising the technology and consolidating usage figures

From a very early stage, OA statistics worked closely with repository operators who already calculated their own usage figures in order to develop a prototype for a reliable service. Practical experience produced a number of improvement opportunities. In order to minimise the amount of data to be transferred, the OAS data provider supplied statistics where irrelevant elements such as Cascading Style Sheets (CSS) and JavaScript elements also present in Apache log files were filtered out. This helped to reduce the size of the database and the amount of data to be transferred. Special robot tables were added to the OAS service provider's database, thus allowing filtered entries to be checked and new robots to be identified based on their entries in the "robots.txt" file (see section 6.5). The repository operators also used this as an opportunity to analyse and compare log files they had already evaluated using the OA statistics software. Any differences between their own usage figures and those returned by OA statistics were traced back to the difference in filtering automated hits by non-identical robot lists.

#### **Excursus: Report from pilot partner EconStor**

From autumn 2009, we collected usage figures using a DSpace add-on based on the statistics plug-in provided by the University of Minho in Portugal. Thanks to our own additional add-ons, this DSpace add-on provided figures that were accurate enough but did not match any known international standards or standards used by German repositories, meaning that it was impossible to compare figures.

While looking for a standard we eventually came across the de-facto standards LogEc provided by RePEc and the COUNTER Code of Practice (COUNTER). LogEc was our standard of choice for two reasons: firstly, EconStor already receives LogEc-compatible usage data as a RePEc data provider for documents that can be downloaded via RePEc's many services; secondly, RePEc is a key player in terms of scientific information.

---

<sup>28</sup> The COUNTER XML specification is available by visiting <http://www.niso.org/schemas/sushi>.

At that stage we were currently in the midst of the DINI certification process, during which we came across the DINI OA statistics project while looking for a RePEc-independent implementation of this process. The final workshop of the first phase of the OA statistics project in January 2010 really aroused our interest to take part in the project as a data provider, especially since both LogEc and COUNTER were to be supported.

The upshot of this project for us was that we were initially able to work without LogEc and that the figures calculated according to COUNTER rules were not fundamentally different from the ones we had calculated ourselves. There was however a number of shortcomings in terms of comparability with the COUNTER figures. These shortcomings derive from the fact that the intervals and criteria, according to which the list of robots and list of federated and automated search engines are updated, are not clear and merely documented as a minimum requirement. This means that each COUNTER user is free to choose the extent to which its repositories usage events are being filtered. The same applies to the paragraph identifying abnormal spikes in usage, which states that “COUNTER does not prescribe a course of action once abnormal spikes in usage have been identified; this is left to the discretion of the customer and vendor.”<sup>29</sup> All of these points constitute arguments in favour of a central service provider as laid out in the OA statistics project since such a provider can guarantee comparability of figures vis-à-vis its data providers. This does not however enable figures from other providers to be compared with usage figures based on COUNTER. We see this as a major problem for EconStor as we set up our own local and therefore independent COUNTER implementation to perform reciprocal OA statistics checks.

Within this context, we also identified two further problems when directly comparing figures for which COUNTER does not provide a precise specification:

1. The time interval of calculation and evaluation: There is no stipulation as to whether the time zone is UTC/GMT or the local time zone including winter and summer time. This is not critical from a statistical point of view, but it does represent an obstacle if two service providers use different time zones or deal with time zones differently when comparing figures for a certain period of time.
2. The behaviour of the 10-second or 30-second rule when exceeding the interval used to analyse the log file (daily, weekly or monthly), e.g. service provider A analyses the log file daily. The log file therefore records a single docu-

---

<sup>29</sup> COUNTER (2012), p. 27.

ment hit from a single IP address at 23h59m56s and another at 0h0m04s the next day. This means that one hit is recorded for each day. Provider B, on the other hand, only analyses the log file once a month and therefore only logs one hit.

In order to ensure that counting on our server is as consistent as possible, we decided to go with our local implementation. We also started to check the usage events on a monthly basis to identify new robots and conspicuous sub-networks and, if necessary, to update our local block lists and robot lists accordingly.

In spite of the problems described above, the positive collaboration with OA statistics meant that we were able to check the respective implementations and optimise them in such a way that we are happy to help maintain the COAR working group's standardised robot list initiated by OA statistics with the aim of achieving improved comparability of usage figures with other repositories.

#### 4.4 Data protection

Log files containing information about the request are used to analyse access to repositories. The log file records the requested document, access results, and information about the requester including their IP address. Requesters include search engines or real people who would like to read the publication they are requesting. In this case, data protection laws must be observed which impose certain requirements and restrictions in terms of processing personal data. According to the survey carried out by the *Zentrale Datenschutzstelle der baden-württembergischen Universitäten* (ZENDAS – Central Data Protection Office for Universities in the German State of Baden-Württemberg)<sup>30</sup> and commissioned by the OA statistics project, both IP addresses and other data collected at the same time must be considered personal data and are therefore subject to prevailing data protection laws. This therefore means that IP addresses have to be pseudonymised using encryption methods, so IP addresses are salted and resalted once a month before being hashed. If all of the participating repositories use the same salt and hash process, document hits with the same document ID can therefore be analysed by any repository. A common salt is not currently in use, although this is currently being trialled and could be easily implemented in future service developments. The central OA statistics service must not be able to recognise this common salt in order to ensure that it cannot decrypt the data. While data from data suppliers, i.e. repositories, are pseudonymised, the encrypted IP addresses can be viewed

---

<sup>30</sup> <http://www.zendas.de/>.

in anonymised form after they have been received by the central OA statistics service. Once the data has been anonymised, it is no longer of importance from a data protection perspective.<sup>31</sup>

### **.htaccess-protected location for the OAS data provider**

For data protection reasons, third parties must not be able to access the original log data or the log data prepared for OA statistics in OpenURL ContextObjects format. For this reason, the OAS data provider's directory is protected from third-party access using .htaccess.

### **Implementing an opt-out**

For reasons pertaining to data protection and/or media law, repositories must provide users with the option to opt-out from having their behaviour collected and analysed. This could, for example, be a button which saves a cookie on the user's computer that is then read out by the web server and prevents the user's IP address from being logged.

## **4.5 Filtering non-human hits**

Open Access repositories are open to any form of access. A large proportion of hits, often 30% or more, originates from automated computer software. Examples of such software include search engines that crawl websites and add them to their indexes as well as programs that search the web for certain content such as e-mail addresses. These so-called robot hits need to be detected and excluded as they falsify usage statistics. Only some robots can be identified in log files, while all other robots only have vague characteristics, meaning that it is hard to tell with any degree of certainty whether a hit originated from a robot or human user. Robot hits can never be fully eliminated from statistics, which is why it is important to develop mutual criteria that can then be used to filter hits in order to be able to compare statistics with one another.

The robot hit filtering intended for the OA statistics project is based on the *user agent* identification principle. This is the term given to the calling program and is saved in the web server log file with every hit and subsequently transmitted to the service providers in the Open URL ContextObjects. This contains a list of well-known robots and their regular expression which is compared with the *user agent* of the respective hit. If the *user agent* of the hit is listed, the hit is considered to originate from a robot.

---

<sup>31</sup> Cf. ZENDAS (2011) p. 25.

The list<sup>32</sup> available on the COUNTER website was taken and then added to using entries from the freely available robotstxt.org list and Open Source statistics software AWStats before being applied to the OA statistics project. This list should be added to and versioned on a regular basis, and also made freely available to the research community. A general robot list such as the one above was discussed at length during the “Usage Statistics and Beyond” (22/23 April 2013) workshop<sup>33</sup> and will be further concretised within the scope of the COAR Interest Group “Usage Data and Beyond”<sup>34</sup>.

New robots can be detected when they access a file called *robots.txt* stored in the web server’s root directory. Even if this is mandatory for robots, it is not a sufficient condition for a robot hit as users can also open the file. The requests are therefore logged in the service provider and available for manual analysis.

---

<sup>32</sup> <http://www.robotstxt.org/db.html>.

<sup>33</sup> <http://www.dini.de/projekte/oa-statistik/english/the-project/workshops/usage-statistics-and-beyond/>.

<sup>34</sup> <http://www.coar-repositories.org/activities/repository-interoperability/usage-data-and-beyond/>.

## 5 Current developments and requirements

After completing the OA statistics project, it can be seen that there is a growing interest in standardised usage figures for digital publications. Repository users are, for example, interested in hit lists sorted by the number of document hits, while authors also want to know how often their document was requested and/or downloaded. Document server operators, on the other hand, are interested in usage figures to help them optimise their services.

The dissemination of scientific texts online gives rise to a number of new ways to analyse their impact, not just by determining absolute document usage figures, but by determining mentions (or citations) of publications in social media services and scientific information repositories such as online literature databases and research data repositories. Some departments and working groups carry out their work online using the internet as a tool, literature database, communication platform and publication location. Research data is published as raw data with the source codes of research trials made publicly available. Scientists publish interim results of their work in their own blog posts or comment on the work carried out by their colleagues. Literature is collated in online literature databases which, in turn, enable networking and collaboration (e.g. Mendeley or ResearchGate) as a further development of community platforms.

Priem et al. called this new impact measurement opportunity “altmetrics” in their policy statement “altmetrics: a manifesto”<sup>35</sup> and added it as a fourth pillar of impact measurement alongside the three other factors: usage, peer review and citations (cf. figure 2).

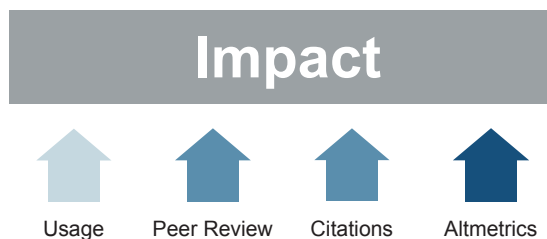


Figure 2: Scientific object impact factors [according to Priem 2011]

Among other things, altmetrics is used to count the number of Twitter mentions and Facebook likes. This means that collection and analysis are no longer limited

---

<sup>35</sup> Cf. Priem (2011).

to the publications themselves as software developments supplied via GitHub or presentations posted on SlideShare can now also be logged. Thanks to their appearance and use in social media, altmetrics are per se readily available and immediately show the visibility and impact of an object online.

There are a number of different providers that collect altmetrics and then process and publish them on various online channels and platforms. The most common altmetrics providers currently available are ImpactStory<sup>36</sup>, Altmetric<sup>37</sup>, PLoS ALM<sup>38</sup> and PLUM Analytics<sup>39</sup>. These providers differ in terms of their target groups, business models and the social media services they use.

The tools offered by these providers can in particular be used to enhance repository documents with alternative metrics. To this end, ImpactStory and Altmetric offer a freely available interface to incorporate the calculated metrics into the repository's document view. All that is required is the document identifier, which is usually a DOI or PubMed ID. Figure 3 shows how the altmetric service was integrated into Göttingen's GoeScholar<sup>40</sup> repository. A DOI is used as an identifier in order to query the provider's API.

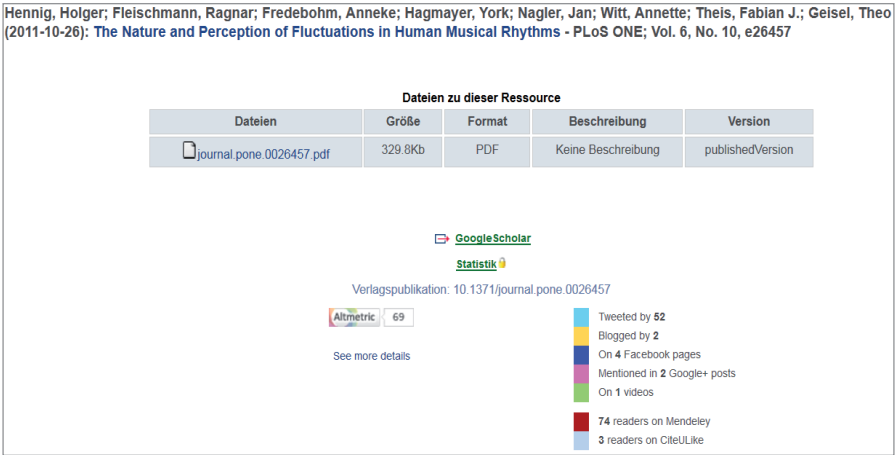


Figure 3: Integrating the altmetric service into Göttingen's GoeScholar repository

<sup>36</sup> <http://impactstory.org/>.

<sup>37</sup> <http://altmetric.com/>.

<sup>38</sup> <http://article-level-metrics.plos.org/>.

<sup>39</sup> <http://www.plumanalytics.com/>.

<sup>40</sup> <http://goedoc.uni-goettingen.de/goescholar>.

The figures calculated by altmetrics still have to be treated with caution as the number of mentions for the same document vary greatly from provider to provider. An example here would be the Altmetric document shown in figure 3 which received 50 mentions on Twitter but only 4 on PLoS ALM and 31 on ImpactStory. It is still unclear as to why these figures are so different, but it does show that standards are required to collect and calculate altmetrics. The NISO (National Information Standards Organisation) has acknowledged this need and is currently investigating the requirements for establishing a standard in a project called the NISO Alternative Assessment Metrics (Altmetrics) Project<sup>41</sup> that is being sponsored for two years. The lack of reliability in terms of data and other needs were also discussed during the altmetrics breakout session<sup>42</sup> at the CERN Workshop on Innovations in Scholarly Communication (OAI8) in Geneva in June 2013. The results of this discussion were recorded in a mindmap<sup>43</sup>: Data storage and management were major discussion topics along with reliability of service, supplied values, document and author identification, and standardisation.

The interaction of the impact factors shown in figure 3 provide a clear overview of a document's impact. The community generally refers to this approach as article-level metrics (ALM), which are measured in various dimensions such as usage (e.g. PDF downloads, HTML downloads), citations (e.g. Web of Science, Scopus), social networks (e.g. Mendeley, CiteULike, Twitter) and mentions in blogs and other online media (e.g. nature blogs, Wikipedia).<sup>44</sup> The Public Library of Science provides an example where ALM is offered for every article (cf. figure 4).

---

<sup>41</sup> [http://www.niso.org/topics/tl/altmetrics\\_initiative/](http://www.niso.org/topics/tl/altmetrics_initiative/).

<sup>42</sup> <https://indico.cern.ch/contributionDisplay.py?sessionId=10&contribId=30&confId=211600>.

<sup>43</sup> Cf. Herb (2013).

<sup>44</sup> Cf. <http://article-level-metrics.plos.org/alm-info/>.





Figure 4: Article-level metrics using PLoS as an example

Reliability and a lack of standardisation are currently opposing the benefits of the altmetrics method (swift, multidimensional analysis of the impact of a broad object pool of scientific information across various platforms). When it comes to usage statistics, OA statistics came up with a way to solve these challenges. If the altmetrics issues can also be solved, it may lead to general acceptance and a rise in impact of alternative metrics. Wouters & Costas (2012) also see major potential for these methods in their evaluation of the further development of impact measurement commissioned by the SURF Foundation: "Altmetrics are booming and they are starting to be seen as alternatives to more conventional citation measures. Metrics on the number of readers, tags used, bookmarks, comments and threads, blogging, tweets, etc. are starting to be suggested as new tools to assess the impact and influence that researchers have over their colleagues and society at large."<sup>45</sup>

---

<sup>45</sup> Cf. Wouters (2012), p. 42.

## 6 Bibliography

### **Bollen 2006**

Bollen, Johan; Van de Sompel, Herbert: *An Architecture for the Aggregation and Analysis of Scholarly Usage Data*. In: Proceedings of the Joint Conference on Digital Libraries, 2006. p. 298–307. URL: <http://dl.acm.org/citation.cfm?doid=1141753.1141821>; doi:10.1145/1141753.1141821.

### **COUNTER (2012)**

Counting Online Usage of Networked Electronic Resources (COUNTER): *The COUNTER Code of Practice for e-Resources: Release 4*. April 2012. URL: <http://www.projectcounter.org/r4/COPR4.pdf>.

### **Dong 2005**

Dong, Peng; Loh, Marie; Mondry, Adrian: *The “impact factor” revisited*. In: Biomedical Digital Libraries (2005), 2:7. URL: <http://www.bio-diglib.com/content/2/1/7>; doi:10.1186/1742-5581-2-7.

### **Herb 2012**

Herb, Ulrich; Mittelsdorf, Björn: *Nutzungsinformationen elektronischer Publikationssysteme: Anwenderwünsche. Ergebnisse dreier empirischer Erhebungen* (Usage information of electronic publication systems: user requests. Results of three empirical surveys), Saarbrücken, 2012. URL: <http://scidok.sulb.uni-saarland.de/volltexte/2012/5008/>.

### **Herb 2013**

Herb, Ulrich: *Mindmap: Acceptance of AltMetrics*. In: scinoptica blog, 20 June 2013. URL: <http://www.scinoptica.com/pages/topics/mindmap-acceptance-of-altmetrics.php>.

### **Merk 2008**

Merk, Christine; Windisch, Nils K.: *Usage Statistics Review: Final report*. JISC, 24 September 2008. URL: [http://repository.jisc.ac.uk/250/1/Usage\\_Statistics\\_Review\\_Final\\_report.pdf](http://repository.jisc.ac.uk/250/1/Usage_Statistics_Review_Final_report.pdf).

### **OA statistics 2012**

OA statistics: *Machbarkeitsstudie zum nachhaltigen Betrieb des Dienstes Open-Access-Statistik* (Feasibility study on long-term operation of the Open Access statistics service). 2012. URL: <http://www.dini.de/fileadmin/oa-statistik/Machbarkeitsstudie/OAS-Machbarkeitsstudie.pdf>.

**Priem 2011**

Priem, Jason; Taraborelli, Dario; Groth, Paul; Neylon, Cameron: Altmetrics: A manifesto. Version 1.01, 28 September 2011.

URL: <http://altmetrics.org/manifesto>.

**Priem 2012**

Priem, Jason; Hemminger, Bradley M.: *Decoupling the Scholarly Journal*. Frontiers in Computational Neuroscience, 6:19. URL: [http://www.frontiersin.org/Computational\\_Neuroscience/10.3389/fncom.2012.00019/abstract](http://www.frontiersin.org/Computational_Neuroscience/10.3389/fncom.2012.00019/abstract).

**Seglen 1997**

Seglen, Per O.: *Why the impact factor of journals should not be used for evaluating research*. BMJ 1997;314:497.1 URL: <http://www.bmj.com/content/314/7079/497.1>.

**Shirky 2008**

Shirky, Clay: *Here Comes Everybody: The Power of Organizing Without Organizations*. New York, The Penguin Press, 2008. p. 81 – 108.

**Verhaar 2010**

Verhaar, Peter et al.: *KE Usage Statistics Guidelines: Guidelines for the aggregation and exchange of usage data*. Version 1.0. SURF Wiki, 18 May 2010. URL: [http://purl.org/REP/standards/KE Usage Statistics Guidelines](http://purl.org/REP/standards/KE_Usage_Statistics_Guidelines).

**Wouters 2012**

Wouters, Paul; Costas, Rodrigo: *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*. SURFFoundation, 2012. URL: <http://www.surf.nl/binaries/content/assets/surf/en/knowledgebase/2011/Users+narcissism+and+control.pdf>

**ZENDAS 2011**

Zentrale Datenschutzstelle der baden-württembergischen Universitäten (ZENDAS – Central Data Protection Office for Universities in the German State of Baden-Württemberg): *Datenschutzrechtliche Bewertung des Projekts Open-Access-Statistik* (Open Access statistics project review in terms of data protection law). Stuttgart, 2011. URL: [http://www.dini.de/fileadmin/oa-statistik/gutachten/ZENDAS\\_Gutachten\\_2011.pdf](http://www.dini.de/fileadmin/oa-statistik/gutachten/ZENDAS_Gutachten_2011.pdf).

## Index of authors

This document was created within the scope of the DFG-funded Open Access statistics project – Services and standards for internationally comparable usage statistics.

### Authors

Justine Haeberli-Kaul, *Universitätsbibliothek Stuttgart*

(Stuttgart University Library)

Daniel Beucke, *Niedersächsische Staats- und Universitätsbibliothek Göttingen*

(Göttingen State and University Library)

Matthias Hitzler, *Niedersächsische Staats- und Universitätsbibliothek Göttingen*

(Göttingen State and University Library)

Andreas Holtz, *Saarländische Universitäts- und Landesbibliothek*

(Saarland University and State Library)

Julika Mimkes, *Niedersächsische Staats- und Universitätsbibliothek Göttingen*

(Göttingen State and University Library)

Wolfgang Riese, *Zentralbibliothek für Wirtschaftswissenschaften*

(German National Library of Economics)

Ulrich Herb, *Saarländische Universitäts- und Landesbibliothek*

(Saarland University and State Library)

Marco Recke, *Verbundzentrale des GBV*

(head office of the GBV Common Library Network)

Birgit Schmidt, *Niedersächsische Staats- und Universitätsbibliothek Göttingen*

(Göttingen State and University Library)

Matthias Schulze, *Universitätsbibliothek Stuttgart*

(Stuttgart University Library)

Sabine Henneberger, *Humboldt-Universität zu Berlin*

(Humboldt-Universität zu Berlin)

Bastian Stemmer, *Zentrale Datenschutzstelle der baden-württembergischen*

*Universitäten* (Central Data Protection Office for Universities in the German State of Baden-Württemberg)

Translated into English: Andrew Rennison





## **Imprint**

DINI – Deutsche Initiative für Netzwerkinformation e.V.

DINI-Geschäftsstelle

c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen

Platz der Göttinger Sieben 1

D-37073 Göttingen

Tel.: +49 551 39-33857

Fax: +49 551 39-5222

E-Mail: [gs@dini.de](mailto:gs@dini.de)

[www.dini.de](http://www.dini.de)